Supporting information

High-fidelity correction of hereditary tyrosinemia in rats through Cas9-nickase-mediated genome editing

Yanjiao Shao[1,†], Liren Wang[1,†], Nana Guo[1], Shengfei Wang[1], Lei Yang[1], Yajing Li[1], Mingsong Wang[1], Shuming Yin[1], Honghui Han[2], Li Zeng[1,2], Ludi Zhang[3], Lijian Hui[3], Qiurong Ding[4], Jiqin Zhang[1], Hongquan Geng[5], Mingyao Liu[1, 6,*] and Dali Li[1,*]

Table of Contents

A **EMX1-sgRNA1+Cas9**

```
            Target             PAM
      GAGTCCGAGCAGAAGAAGAAGGG
GGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGG       WT
GGCCTGAGTCCGAGCAG---AAGAAGGGCTCCCATCACATCAACCGGTGG       -3bp   x 5
GGCCTGAGTCCGAGC------AGAAGGGCTCCCATCACATCAACCGGTGG       -6bp   x 8
GGCCTGAGTCCGAGCAGAAGA------GCTCCCATCACATCAACCGGTGG       -6bp   x 2
GGCCTGAGTCCG---------AGAAGGGCTCCCATCACATCAACCGGTGG       -9bp   x 1
GGCCTG---------AG---CAGAAGGGCTCCCATCACATCAACCGGTGG       -12bp  x 1
GGCCTGAG-------------AGAAGGGCTCCCATCACATCAACCGGTGG       -13bp  x 1
GGCCT-----------------GAAGGGCTCCCATCACATCAACCGGTGG       -17bp  x 1
GGCCTGAGTCCGAGCAG----AAGAA------------------CGGTGG       -22bp  x 1
----------------------GAAGGGCTCCCATCACATCAACCGGTGG       -26bp  x 1
-----------------------------TCCCATCACATCAACCGGTGG       -100bp x 1
GGCCTGAGTCCGAGCAGAAGAAAGAAGGGCTCCCATCACATCAACCGGTGG      +1bp   x 20
GGCCTGAGTCCGAGCAGAAGAATTTACAGTCAAATTAATTCTAATTATCTCTCTAACAGCCTTGTATCCATCA
CATCAACCGGTGG                                            +36bp  x 1
```

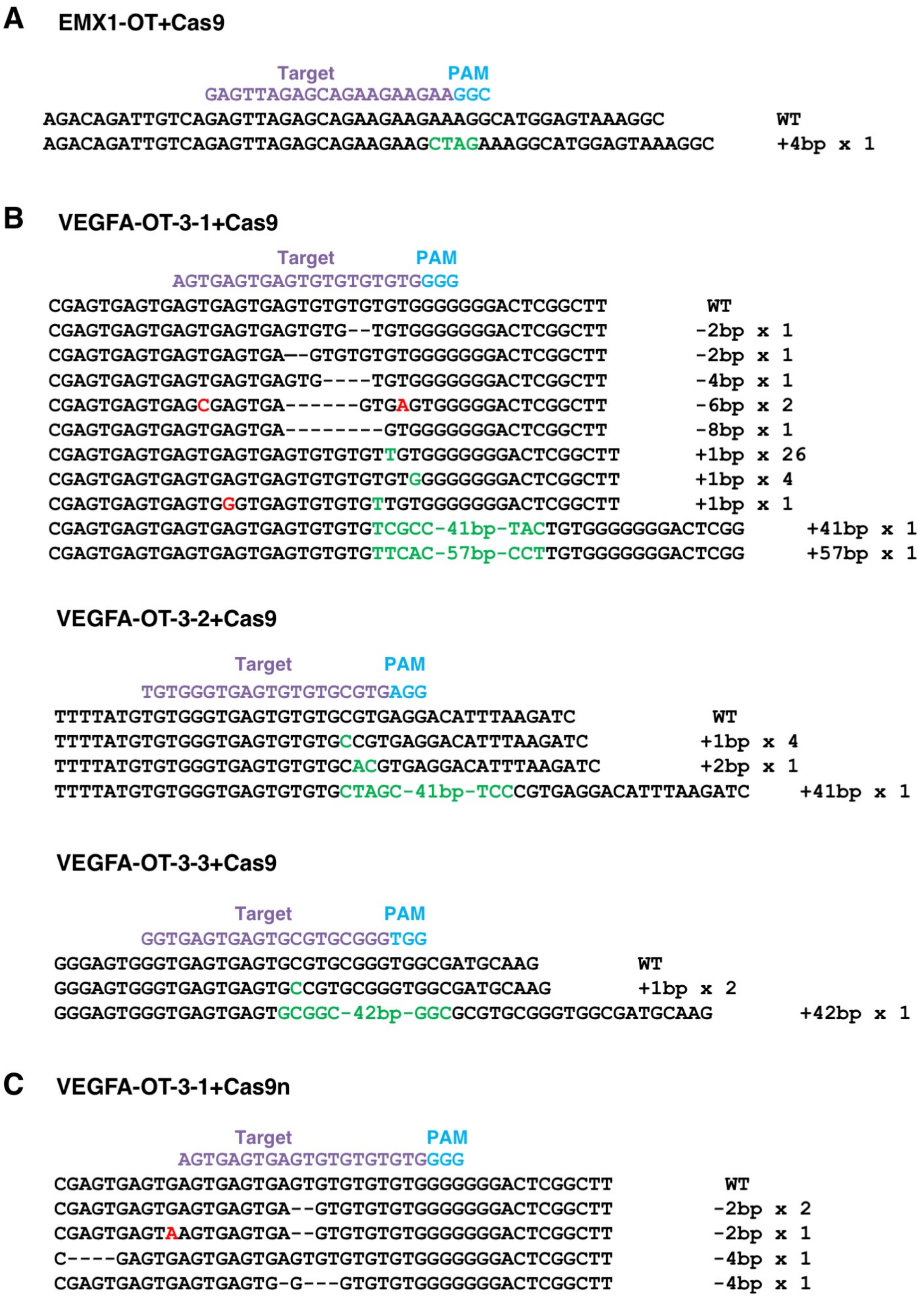
B **EMX1-sgRNA1+Cas9n**

```
            Target             PAM
      GAGTCCGAGCAGAAGAAGAAGGG
GGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGG       WT
GGCCTGAGTCCGAACTTAAGAAGAAGGGCTCCCATCACATCAACCGGTGG       x 1
```

C **VEGFA-sgRNA3+Cas9**

```
             Target            PAM
      GGGTGAGTGAGTGTGTGCGTGTGG
TGTGTGGGTGAGTGAGTGTGTGCGTGTGGGGTTGAGGGCGTTGG             WT
TGTGTGGGTGAGTGAGTGT----GTGTGGGGTTGAGGGCGTTGG             -4bp x 1
TGTGTGGGTGAGTGAGT------GTGTGGGGTTGAGGGCGTTGG             -6bp x 6
TGTGTGGGTGAGCGA--------GTGTGGGGTTGAGGGCGTTGG             -8bp x 9
TGTGTGGGTGAGTGAGTGTG---------------------TGG            -21bp x 1
TGTGTGGG------------------------TGAGGGCGTTGG            -24bp x 1
TGTGTGGG------------------------TGAGTGCGTTGG            -24bp x 1
TGTGTGGGTGAGTGAGTGTGTGCCGTGTGGGGTTGGGGGCGTTGG            +1bp x 14
TGTGTGGGTGAGTGAGTGTGTGCCCGTGTGGGGTTGAGGGCGTTGG           +2bp x 1
TGTGTGGGTGAGTGAGTGTGTGTCTGTGTGGGGTTGAGGGCGTTGG           +2bp x 1
TGTGTGGGTGAGTGAGTGTGTGCCTCGTGTGGGGTTGAGGGCGTTGG          +3bp x 1
TGTGTGGGTGAGTGAGTGTGTGCGACGCCTACCTGAACGCCGTCGTGTGGGGTTGAGGGCGTTGG +18bp x 1
TGTGTGGGTGAGTGAGTGTGTGCGTCAGAGAC-18bp-CTCACCGTGTGGGGTTGAGGGCGTTGG +18bp x 1
TGTGTGGGTGAGTGAGTGTGTGCTACGCC—108bp-TCGCCCTCGTGTGGGGTTGAGGGCGTTGG +108bpx 1
```

Fig. S1. Indels generated by Cas9 or Cas9n in on-target sites in HEK293T cells. Indels detected in the on-target sites of (a) the EMX1 locus by Cas9, (b) the EMX1 locus by Cas9n, (c) the VEGFA locus by Cas9. The occurrence frequencies of each on-target indel clone are listed at the right.



```
                  GAGTTAGAGCAGAAGAAGAAGGC
AGACAGATTGTCAGAGTTAGAGCAGAAGAAGAAAGGCATGGAGTAAAGGC          WT
AGACAGATTGTCAGAGTTAGAGCAGAAGAAGCTAGAAAGGCATGGAGTAAAGGC      +4bp x 1
```

```
             AGTGAGTGAGTGTGTGTGTGGGG
CGAGTGAGTGAGTGAGTGAGTGTGTGTGTGGGGGGGACTCGGCTT         WT
CGAGTGAGTGAGTGAGTGAGTGTG--TGTGGGGGGGACTCGGCTT        -2bp x 1
CGAGTGAGTGAGTGAGTGA—-GTGTGTGTGGGGGGGACTCGGCTT        -2bp x 1
CGAGTGAGTGAGTGAGTGAGTG----TGTGGGGGGGACTCGGCTT        -4bp x 1
CGAGTGAGTGAGCGAGTGA------GTGAGTGGGGGACTCGGCTT        -6bp x 2
CGAGTGAGTGAGTGAGTGA--------GTGGGGGGGACTCGGCTT        -8bp x 1
CGAGTGAGTGAGTGAGTGAGTGTGTGTTGTGGGGGGGACTCGGCTT       +1bp x 26
CGAGTGAGTGAGTGAGTGAGTGTGTGTGGGGGGGGACTCGGCTT         +1bp x 4
CGAGTGAGTGAGTGGGTGAGTGTGTGTTGTGGGGGGGACTCGGCTT       +1bp x 1
CGAGTGAGTGAGTGAGTGAGTGTGTGTCGCC-41bp-TACTGTGGGGGGGACTCGG      +41bp x 1
CGAGTGAGTGAGTGAGTGAGTGTGTGTTCAC-57bp-CCTTGTGGGGGGGACTCGG      +57bp x 1
```

```
         TGTGGGTGAGTGTGTGCGTGAGG
TTTTATGTGTGGGTGAGTGTGTGCGTGAGGACATTTAAGATC             WT
TTTTATGTGTGGGTGAGTGTGTGCCGTGAGGACATTTAAGATC           +1bp x 4
TTTTATGTGTGGGTGAGTGTGTGCACGTGAGGACATTTAAGATC          +2bp x 1
TTTTATGTGTGGGTGAGTGTGTGCTAGC-41bp-TCCCGTGAGGACATTTAAGATC    +41bp x 1
```

```
         GGTGAGTGAGTGCGTGCGGGTGG
GGGAGTGGGTGAGTGAGTGCGTGCGGGTGGCGATGCAAG         WT
GGGAGTGGGTGAGTGAGTGCCGTGCGGGTGGCGATGCAAG        +1bp x 2
GGGAGTGGGTGAGTGAGTGCGGC-42bp-GGCGCGTGCGGGTGGCGATGCAAG       +42bp x 1
```

```
          AGTGAGTGAGTGTGTGTGTGGGG
CGAGTGAGTGAGTGAGTGAGTGTGTGTGTGGGGGGGACTCGGCTT          WT
CGAGTGAGTGAGTGAGTGA--GTGTGTGTGGGGGGGACTCGGCTT         -2bp x 2
CGAGTGAGTAAGTGAGTGA--GTGTGTGTGGGGGGGACTCGGCTT         -2bp x 1
C----GAGTGAGTGAGTGAGTGTGTGTGTGGGGGGGACTCGGCTT         -4bp x 1
CGAGTGAGTGAGTGAGTG-G---GTGTGTGGGGGGGACTCGGCTT         -4bp x 1
```

Fig. S2. Off-target events of Cas9 or Cas9n in HEK293T cells. Indels detected in off-target sites of (a) the EMX1 locus by Cas9, (b) the VEGFA locus by Cas9 and (c) the VEGFA locus by Cas9n. The occurrence frequencies of each off-target indel are listed on the right. Base substitutions and insertions are labeled in red and green respectively. The occurrence frequencies of each off-target indel are listed on the right.

***Fah locus* +AdV-Cas9+AdV-HDR**

```
TGTCACTCTTCTCCTAGCCAAAGCCACGGATTGGTGTGACCAGATCTTGGACCTGAGTGTCATTAAACAC FahΔ10/Δ10
TGTCACTCTTCTCCTAGCCAAAGCCACGGATTGGAGTGACCAGATCTTGGACCTGAGTGTCATTAAACAC +1bp
TGTCACTCTTCTCCTAGCCAAAGCCACGGATTGGTGTGACCAGA--TTGGACCTGAGTGTCATTAAACAC -2bp
TGTCACTCTTCTCCTAGCCAAAGCCACGGATTGGTGT------------GACCAGAGTGTCATTAAACAC -12bp
TGTCACTCTTCTCCTAGCCAAAGCCACGGATTGG-------------TGGACCTGAGTGTCATTAAACAC -13bp
TGTCACTCTTCTCCTAGCCAAAGCCACGGATTGGT----------------CCTGAGTGTCATTAAACAC —16bp
TGTCACTCTTCTCCTAGCCAAAG----------------------------CCTGAGTGTCATTAAACAC -28bp
TGTCACTCTTCTCCTAGCCAAAG----------------------------C---------------CAC -43bp
TGTCAC-----------------------------------------TGGACCTGAGTGTCATTAAACAC -41bp
TGTCA--------------------------------------------------AGTGTCATTAAACAC -50bp
TGTCACTCTTCTCCTAGCCAAAGCCACGGAAGCTGA.(total 36bp).ATTGGTGTGTCATTAAACAC +36bp
```

Fig. S3. Indels generated by AdV-Cas9. Sanger sequencing of 98 clones of the amplicons from the Adv-Cas9 treated rat revealed indels caused by Cas9-induced NHEJ. Indel sizes are listed on the right. The target sequence is indicated by the purple characters and the PAM sequence is in blue.

Table S1 Summary of animal groups

Injected adenovirus[+]	**NTBC withdrawal**	**Surviving**[§]**/total animals**
Cas9 + HDR template	Yes	3/4
Cas9n + HDR template	Yes	3/4
HDR template	Yes	0/4
PBS	Yes	0/4

[+] Animals received total $2x10^{10}$ virus in 800 µL PBS, including 1x1E10 Cas9 (WT & nickase) and 1x1E10 HDR template.

[§] NTBC free survival for at least 3 months.

Table S2 On-target /off-targt sites and primers used in this study.

On-target and Off-target sites			
Target ID	Target sequence (5'-3')+pam	Target ID	Target sequence (5'-3')+pam
Fah$^{+/+}$ on-target	CCACGGATTGGTGTGGCCATCGG	Fah_off_2	GGACTGGGGTAACCAGATCTTAG
Fah$^{\Delta10/\Delta10}$ on-target	GGATTGGTGTGACCAGATCTTGG	Fah_off_3	GGATGGGAGTGACCAGCTCTCAG
EMX1 on-target	GAGTCCGAGCAGAAGAAGAAGGG	Fah_off_4	TCACTGGTTTGACCAGATCTGGG
EMX1 OTS	GAGTTAGAGCAGAAGAAGAAAGG	Fah_off_5	GTACTGGCATGACCAGATCTGGG
VEGFA on-target	GGTGAGTGAGTGTGTGCGTGTGG	Fah_off_6	GGAATTGTGTGACCAGATCAGGG
VEGFA OTS-3-1	AGTGAGTGAGTGTGTGTGTGGGG	Fah_off_7	GTCCTGGTGTCACCAGATCTGAG
VEGFA OTS-3-2	TGTGGGTGAGTGTGTGCGTGAGG	Fah_off_8	TCAATGGTGTGGCCAGATCTCAG
VEGFA OTS-3-3	GGTGAGTGAGTGCGTGCGGGTGG	Fah_off_9	GCCTCGGTGTGGCCAGATCTTGG
Fah_off_1	GCAATGGTGGGACCAGATCTGGG	Fah_off_10	GCCTGGGTGTGCCCAGATCTGAG
Primers for PCR amplification			
Primer ID	**Primer sequence**	**Primer ID**	**primer sequence**
Fah PCR-F	CTGCTGCATTTAAGCTACCAC	Fah PCR-R	CATAGAAGCCAGGATGAGTGT
EMX1-F	CTCTTCCTGCCCTGCCATCC	EMX1-R	CCAGCCCATTGCTTGTCCCT
EMX1 OT-F	ATACAATGGGAAGGACAGC	EMX1 OT-R	TGTGGAGGCAATAAAGTGA
VEGFA 3-F	CAGATGGCACATTGTCAGAGGG	VEGFA 3-R	GAGGGAGCAGGAAAGTGAGGTTA
VEGFA OT-3-1-F	GTCGCCTGGATGCGGTTTC	VEGFA OT-3-1-R	CGTTATCGCTCATTTCCTACGG
VEGFA OT-3-2-F	TGTATCCCTGGATTTCGTAT	VEGFA OT-3-2-R	ATCTAATGTATGGCATGGTG
VEGFA OT-3-3-F	CTGCATTTCTTTGGCGTCTC	VEGFA OT-3-3-R	AGCACAGCCCTTACTCCTCC
rFah-OT1-F	ATTCAGAGTAGCATTCCCCACAGCAGC	rFah-OT1-R	TCTCGAGAGGGGAGTTTCAGGGACCAC
rFah-OT2-F	TATAGCCTAGAGTTCCCTCATTTTGTCC	rFah-OT2-R	ATAGAGGCACTGTCCTATTTATCCATTCG
rFah-OT3-F	CCTATCCTGAAGGAAGATGAGAAATGA	rFah-OT3-R	GGCTCTGACTGTTTTCCCTTTAGAGTT
rFah-OT4-F	AGGCGAAGGGTCATCTCCTGCCTCTGC	rFah-OT4-R	TAATCTTAAGTGGGTAAGTCTAAGAATACA
rFah-OT5-F	CAGGACGTCTGTCATGGGACTCCTCGTC	rFah-OT5-R	GTACTGACTACAGGCAAATCCCGTTC
rFah-OT6-F	ATTACTCGAAGAAAGCCAAGAATGACA	rFah-OT6-R	TCCGGAGAACCATCGACAGTTGAGGAC
rFah-OT7-F	CGCTCATTAGCAGGGCTCTGGTCAGGA	rFah-OT7-R	GAGATTCCGCCGAGGCTCAGGGAAACA
rFah-OT8-F	GAATTCGTTCCCAATGGGATAGTTAAG	rFah-OT8-R	CTGAAGCTTTTGCCTCACTCAGTTCAC
rFah-OT9-F	TAATGCGCAGTTCTTACCCATAAGCCATC	rFah-OT9-R	CGGCTATGCCAAGACCTTTAGACTCCCT
rFah-OT10-F	TCCGCGAAAGGAGGTTTGGGTAAAGACA	rFah-OT10-R	AGCGATAGTTTGAAGCCTACATACGAG
IFN-β1 PCR-F	TGCCATTCAAGTGATGCTCC	IFN-β1 PCR-R	CACCCAAGTCAATCTTTCCTCT
IL6 PCR-F	GTCTTCTGGAGTTCCGTTTC	IL6 PCR-R	GATGGTCTTGGTCCTTAGCC
IL10 PCR-F	CAGGACTTTAAGGGTTACTTGG	IL10 PCR-R	CATTCTTCACCTGCTCCACT
β-actin PCR-F	GTACGCCAACACAGTGCTG	β-actin PCR-R	CGTCATACTCCTGCTTGCTG

Table S3 Off-target analyses by deep sequencing. Potential off-target sites predicted and scored by CRISPR Design for the on target sgRNA. Mismatched base pairs are labeled in red.

ID	Score	Spacer + PAM	Coordinate	Mismatches	Indel Adv-Cas9	Indel Adv-CAs9n	Indel PBS
Fah_sgRNA		**GGATTGGTGTGACCAGATCTTGG**					
Fah_off_1	2.5	**GCAATGGTGGGACCAGATCTGGG**	chr1:+283240856	3	0.0039	0.0039	0.0033
Fah_off_2	1.4	**GGACTGGGGTAACCAGATCTTAG**	chr19:-10419394	3	0.0019	0.0018	0.0016
Fah_off_3	1.1	**GGATGGGAGTGACCAGCTCTCAG**	chr8:-80645234	3	0.0013	0.0015	0.0012
Fah_off_4	0.9	**TCACTGGTTTGACCAGATCTGGG**	chr2:-137494772	4	0.001	0.001	0.0007
Fah_off_5	0.8	**GTACTGGCATGACCAGATCTGGG**	chr1:-124157008	4	0.0025	0.0022	0.0021
Fah_off_6	0.8	**GGAATTGTGTGACCAGATCAGGG**	chr4:+228450195	3	0.001	0.0008	0.002
Fah_off_7	0.8	**GTCCTGGTGTCACCAGATCTGAG**	chr15:-38501543	4	0.0017	0.0015	0.0015
Fah_off_8	0.7	**TCAATGGTGTGGCCAGATCTCAG**	chr7:+78415288	4	0.0013	0.0013	0.0009
Fah_off_9	0.7	**GCCTCGGTGTGGCCAGATCTTGG**	chr16:+7326951	4	0.0014	0.0014	0.0011
Fah_off_10	0.7	**GCCTGGGTGTGCCCAGATCTGAG**	chr18:-30124498	4	0.002	0.0013	0.0015

Supplemental sequence HDR template sequence (5’-3’)

ACACTATTATGCCACTATTACAGCATTGTGGAGATCTTGCTATCCCAGTGATTGTTGGGATACAT
AGGCATTAGAGCTGGGTAAGGCTACTCATAGCTCTCCTCGCTTGGCAACTTATAAAGCACTAA
AATGACTTAATACAATTGGGAGGCCTTGTTCACACATGACTAGGCAGGTAATGCCTGTGATCA
GGTAACAGATGACTTAGTGAGAGGTTCCTTCTTGCCTTCAGCAAGGCATAGGATGATTGCCTC
TGAGTCTTTCCTGAACACCACCTGAGTGGACCCCCCAACTTAAGTAAATTCAGCTTGCTGATC
TTTGTGGCTGAATGAAGGCCCAGGGAATTACAGAGCTGAGAACACACAGGTTCTCCAGGGAA
GCTTCTAAGGTTTATGTGGCCTTAGTCTACCCCTAATCCCTGGGGGCTTCCAGGCAGCTACTTC
TCTTAGCCCACTGAAAGGAGCACTGGAGACAGACACAGCCAAGGATTCTACATCTCACCTCT
CCTGGATCTCCTAGAATATGGCAGAACTGCTGCTTCAGGTGCCAAGGAAGCCTAGCTTAGGAC
AGAGTTCCAGGAGGTGAAATCTTCCAACATGCTCAGTGCTGTGAGGTCAGAGACCAGCCTTG
ACAAGATATTGGGCCTCTGTGTTCATTTGAATCCATGCAGAGCTCTTACAGTGCATAAAAGAA
ATCCATTCATTTGGTCTGAGCACTGTGCCTTCTGCAGTTCCCTACGGCCTGAGCTCTGCTGCAT
TTAAGCTACCACAGAGCTGATACTCCAAGTTACTACATGTGTAACACTCACTGCTCTGCTTCAT
AAGAGAATGGCCCTGAGCAAAGCCCTGGGCCGAGCATGAAAAGAGGTTCGGGTTCTTGGTA
AGGTGATGAGCAACTCATAACTAGAGATCATGGCTCTGCTTTCTTTCTTGTCCTGGATCCCGTC
TCAATGGAATCCATGAGTCCCTAAGTAACTGTGCTGTGACTGTCAGAACTGACTGTCACTCTT
CTCCTAGCCAAAGCCAAGGATCGGAGTCACGTGGACCTGAGTGTCATTAAACACCTCTTTACC
GGACCTGTCCTCTCCAAACATCAGCATGTCTTCGATGAG GTAGGACATTGTGTCATGGACTGT
CTCTGTCTCACCACCTACTACAAATGACATAGCATCTTCATAGTCACCAGTTATTCCCGGATGA
GGAGGAGCTGTTCTGAGACACACAGGTGGCATGAGGTAGCCTCGACACCTTAGGGCAGCTTC
TCAATAGAATGTGAGCTCCTGAGTCCAGTGGATCTCTCCAAGTTGCCTATGAGGTGAGACTGT
TGTGTTAGCCTGGTTCCTTTATATACAAGGATGTTCATGCTTTGGTCTTTGGAGGCTCAGGAAA
ACCTTGAGAAAAGGAGCCATGAGCCTTGCAGTCTAGATTTTAGGGTAAAGCTTGAGACACTC
ATCCTGGCTTCTATGTTCATGAATGATGACTGCCCAGTACTTCAAGGGTACAGTCCTGGGGCAT
CCAACCTTGCTTGTGTGTCTGTGTAGGATGCCTACCCTGTCATCCAGGTATGCAGTTGTGCCTG
GGGGCAGCCAGCTTGTCTACCTCGTCTGTGTGGGATGGATGTAACTTTC